CLAIMS

What is claimed is:

1   1.   A method for permitting a subscriber to perform an action available on a

2   communications network using a spoken utterance, comprising:

3   maintaining a system state database comprising a tree structure having a plurality

4   of nodes, each respective node of said plurality of nodes representing a particular system state of

5   a plurality of possible system states and being associated with a predetermined node-specific

6   grammar for the respective node;

7   awaiting from the subscriber a spoken utterance at the particular system state;

8   recognizing the spoken utterance by comparing the spoken utterance to the

9   predetermined grammar for the respective node for correspondence to the particular system state;

10 and

11   performing an action at the network represented by the spoken utterance if the

12 spoken utterance has been recognized as the predetermined grammar for the respective node.

1   2.   The method of claim 1, further comprising, after recognizing the spoken

2   utterance, converting the spoken utterance to electronically-readable data having a format

3   recognizable by one of the network, and transmitting the converted data to the respective one of

4   the network.

1   3.   The method of claim 1, wherein the spoken utterance comprises a

2   command to access one of an available feature of a plurality of features available on the network

3   and a spoken menu of the available features.

1   4.   The method of claim 3, wherein the feature comprises one of a group

2   consisting of call forwarding, hold, conferencing, voice-mail, call back, caller-ID, caller-ID

3   related features and caller-ID related functions.

1   5.   The method of claim 1, wherein the node-specific grammar associated

2   with each respective node comprises at least one of a group consisting of a word descriptive of

3   the action to be performed, a synonym of the word, and a globally-available word available at all

4   of said plural nodes.

1   6.   The method of claim 1, wherein the predetermined grammar for the

2   particular node comprises grammar for multiple languages.

1   7.   The method of claim 6, wherein the spoken utterance of the subscriber is

2   in one of the multiple languages, and the method further comprises the steps of: determining the

3   one of the multiple languages of the spoken utterance of the subscriber; and communicating via

4   the network with the subscriber via a text-to-speech translator that translates in the determined

5   one language of the subscriber.

39

1   8.   The method of claim 1, further comprising determining a particular

2   template to use for speech recognition from a plurality of predefined voice pattern templates,

3   wherein the particular template comprises a subset of the predetermined grammar for the

4   respective node, and wherein the step of recognizing the spoken utterance comprises comparing

5   the spoken utterance to the predetermine subset of the predetermined grammar for the respective

6   node.

1   9.   The method of claim 8, wherein the plurality of predefined voice pattern

2   templates comprises independent templates for males, females, and children.

1   10.   The method of claim 1, further comprising the step of prompting the

2   subscriber to issue the spoken utterance using one of a group consisting of a spoken menu

3   generated by a text to speech translator, a recorded announcement of a menu, and a synthesized

4   announcement of the menu.

1   11.   The method of claim 1, further comprising the steps of: transmitting, by

2   the network, a signal to the subscriber in a data format not audibly recognizable by the

3   subscriber; and converting the transmitted signal to an audible message recognizable to the

4   subscriber using one of a text to speech translator, a recording of speech, and a speech

5   synthesizer.

1   12.   The method of claim 11, wherein the signal transmitted by the network to

2   the subscriber comprises one of the group consisting of an ADSI signal and a DTMF signal.

1   13.   The method of claim 1, wherein the action performed comprises

2   transmitting, by the network, of a signal to a second network.

1   14.   The method of claim 1, wherein the method is performed by a speech

2   recognition system, and the method further comprises the step of providing to the subscriber an

3   ability to operatively toggle on and off the speech recognition system.

1   15.   The method of claim 1, wherein the system state database is located on a

2   speech processing unit coupled to the network through one of the group consisting a local

3   communications office equipment, the Internet, a computer, a mobile phone, a headset, a

4   handset, a base station, a set-top box, a personal digital assistant, an appliance, and a remote

5   control, and wherein said step of comparing the spoken utterance is performed at the location of

6   the system state database.

1   16.   The method of claim 1, wherein the plurality of possible system states

2   comprises a plurality of possible steps in a call flow and an "always connected" state in which a

3   feature may be accessed even when a call is not in progress.

1   17. The method of claim 1, further comprising: inputting a key input, and

2 wherein the step of performing the action comprises performing the action in accordance with the

3 spoken utterance and the key input.

1   18. A communications system providing speech recognition functionality to a

2 network, comprising:

3   a device coupled to the network and into which an utterance may be spoken by a

4 user,

5   a system state database accessible to the network and defining a tree structure

6 having a plurality of nodes, each respective node of said plural nodes representing a particular

7 step of a plurality of possible system states and being associated with a predetermined node-

8 specific grammar for the respective node;

9   means for interpreting the user-spoken utterance;

10   means for comparing the interpreted spoken utterance to the predetermined

11 grammar for the respective node corresponding to the particular system state to recognize the

12 spoken utterance as corresponding to the predetermined grammar associated with the respective

13 node; and

14   means for performing an action represented by the spoken utterance at the

15 network if the spoken utterance has been recognized as corresponding to the predetermined

16 grammar associated with the respective node.

1   19.   The communications system of claim 18, wherein the spoken utterance

2   comprises one of a group consisting of a command to access a feature available at the network,

3   and a spoken menu of available features at the network.

1   20.   The communications system of claim 18, wherein the spoken utterance

2   comprises a command to access a feature available at the network, the feature comprising one of

3   a group consisting of call forwarding, hold, conferencing, voice-mail, call back, and caller-ID.

1   21.   The communications system of claim 18, wherein said interpreting means

2   comprises an utterance verification engine.

1   22.   The communications system of claim 18, wherein said comparing means

2   comprises a reference database which comprises the predetermined node-specific grammar

3   associated with each respective node.

1   23.   The communications system of claim 22, wherein the system state and

2   reference databases are both maintained on a speech processing unit coupled to the network

3   through one of a group consisting of a local communications office equipment, the Internet, a

4   computer, a mobile phone, a headset, a handset, a base station, a set-top box, a personal digital

5   assistant, an appliance, and a remote control.

1   24.   The communications system of claim 22, wherein the node-specific

2   grammar associated with each respective node comprises at least one of a group consisting of a

3   word that is descriptive of the action to be performed, a synonym of said at least one word, and a

4   globally-available word available at all of said plural nodes.

1   25.   The communications system of claim 18, wherein the predetermined

2   grammar for the particular node comprises grammar for multiple languages.

1   26.   The communications system of claim 25, further comprising means for

2   determining the language of the spoken utterance of the user, and a text-to-speech translator for

3   translating communications from a network to the user in the determined language of the user.

1   27.   The communications system of claim 18, further comprising means for

2   offering the user a spoken menu of the predetermined grammar available at the respective node

3   in the call flow.

1   28.   The communications system of claim 27, further comprising means for

2   receiving the requested spoken menu and at least a partial text menu of the available features.

1   29.   The communications system of claim 18, further comprising means for

2   transmitting, to the user, a signal in a data format not audibly recognizable by the user, a text to

3   speech translator, and means for converting the transmitted signal to an audible message

4   recognizable to the user using the text to speech translator.

1   30.   The communications system of claim 29, wherein the transmitted signal

2   comprises one of a group consisting of an ADSI signal and a DTMF signal.

1   31.   The communications system of claim 18, wherein the means for

2   performing an action comprises means for transmitting a signal transmitted between networks.

1   32.   The communications system of claim 18, further comprising means for

2   toggling on and off the speech recognition and text-to-speech functionality.

1   33.   The communications system of claim 18, wherein the plurality of possible

2   system states comprises a plurality of possible steps in a call flow and an "always connected" state

3   in which a feature may be accessed even when a call is not in progress.

1   34.   The communications system of claim 18, further comprising: means for

2   inputting a key input, and wherein the means for performing the action comprises performing the

3   action in accordance with the spoken utterance and the key input.